



Aalborg Universitet

AALBORG UNIVERSITY  
DENMARK

## Physical Models for Fast Estimation of Guitar String, Fret and Plucking Position

Hjerrild, Jacob Møller; Willemsen, Silvin; Christensen, Mads Græsbøll

*Published in:*

2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, WASPAA 2019

*DOI (link to publication from Publisher):*

[10.1109/WASPAA.2019.8937157](https://doi.org/10.1109/WASPAA.2019.8937157)

*Publication date:*

2019

*Document Version*

Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*

Hjerrild, J. M., Willemsen, S., & Christensen, M. G. (2019). Physical Models for Fast Estimation of Guitar String, Fret and Plucking Position. In *2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, WASPAA 2019* (pp. 155-159). [8937157] IEEE. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) <https://doi.org/10.1109/WASPAA.2019.8937157>

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### Take down policy

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.

# PHYSICAL MODELS FOR FAST ESTIMATION OF GUITAR STRING, FRET AND PLUCKING POSITION

Jacob Møller Hjerrild,<sup>1,2</sup> Silvin Willemsen,<sup>3</sup> Mads Græsbøll Christensen<sup>2</sup>

<sup>1</sup> TC Electronic, Aarhus, Denmark

<sup>2</sup> Audio Analysis Lab, CREATE, Aalborg University, Denmark

<sup>3</sup> Multisensory Experience Lab, CREATE, Aalborg University, Denmark

jacob.mollerhjerrild@musictribe.com sil@create.aau.dk mgc@create.aau.dk

## ABSTRACT

In this paper, a novel method for analyzing guitar performances is proposed. It is both fast and effective at extracting the activated string, fret, and plucking position from guitar recordings. The method is derived from guitar-string physics and, unlike the state of the art, does not require audio recordings as training data. A maximum a posteriori classifier is proposed for estimating the string and fret based on a simulated model of feature vectors while the plucking position is estimated using estimated inharmonic partials. The method extracts features from audio with a pitch estimator that estimates also the inharmonicity of the string. The string and fret classifier is evaluated on recordings of an electric and acoustic guitar under noisy conditions. The performance is comparable to the state of the art, and the performance is shown to degrade at SNRs below 20 dB. The plucking position estimator is evaluated in a proof-of-concept experiment with sudden changes of string, fret and plucking positions, which shows that these can be estimated accurately. The proposed method operates on individual 40 ms segments and is thus suitable for high-tempo and real-time applications.

**Index Terms**— Physical Modeling, Statistical Signal Processing, Parametric Pitch Estimation, Music Information Retrieval

## 1. INTRODUCTION

The analysis of individual instruments in musical performances has various applications, such as music learning, detailed transcription, artist recognition and extraction of stylistic details. Analysis and synthesis of plucked string instruments have been studied for electric [1] and acoustic guitars [2, 3]. The present paper is concerned with the analysis of guitar string signals for detailed transcription. To date, there are only few papers on extracting information from electric guitar signals, such as the work concerned with multipitch transcription [4] or classifying the types of effects used [5] and estimating decay times of electric guitar tones [6]. Other research involved the extraction of information from plucked string instruments, such as plucking position [7–10] or plucking styles and dynamics for classical guitar [11] and electric bass guitar [12]. Recent papers introduce models of the physical interactions between player and string to make synthesized guitar sound more realistic, such as interactions with the string and guitar pick [13, 14] or the fingers [15], and the fingers with the fretboard [16].

It is well-known that the transducer position and plucking position produce a comb-filtering effect in the guitar signal spectrum [17] and that stringed instruments are inharmonic because of stiffness in the string [18]. Inharmonicity is also related to the

plucking deflection as shown in [19–22]. The inharmonicity has been proven useful for detailed transcription [23], a topic in its infancy, wherein not only notes are identified but also where and how they are played (i.e. string, fret and plucking positions). In [24], a string and fret classification method was proposed using a 10-dimensional feature set and a SVM classifier. In [12], the electric guitar strings were distinguished with a string model comprising a 48-dimensional feature set, wherein the inharmonicity coefficient was selected as one of the most discriminative features.

Large feature sets are prone to overfitting and rarely contribute to meaningful cause and effect findings. On top of this, they also typically results in computationally complex algorithms. To overcome such problems, [23] proposed an inharmonicity and amplitude based method for accurate generation of guitar tablature where the fundamental frequency was assumed to be known, and the main parameter for classification of string and fret was based on counting the number of partials that follow the piano model derived in [25]. Recently, [26] proposed to classify string and fret by modeling each inharmonicity coefficient as a Gaussian distribution. Both methods [23, 26] operate on multiple segments, each in the order of 100 ms, making them unsuitable for high-tempo and real-time applications. To overcome such problems, [27] proposed a fast inharmonic parametric pitch estimation algorithm that operates on a single 40 ms segment to estimate string, fret and the plucking position. The methods presented in [23, 26, 27] require training data in the form of audio recordings to build a model of the guitar, which can be very many, i.e., for 6 strings and 12 frets that is  $K = 72$  classes.

Although [23, 27] showed that a model can be trained from audio recordings captured from only one fret per string, there are three main challenges associated with such a procedure:

- With only a few observations (e.g., 10 per string [27]), it is impossible to conclude anything significant with respect to the covariance structure of each class in the feature space.
- Although training is relatively fast, it is required for every different set of guitar strings.
- There is no information of class dependent distributions in the observed feature space, i.e., covariance structures, except for the 6 classes represented in the training set.

In this paper, we extend the work of [27] by eliminating the training procedure that requires the user to provide recordings and as we shall see, we obtain detailed and class dependent distributions in the feature space. Instead of training a model from audio recordings, we propose to simulate a model of guitar string vibrations based on physical properties of the string, such as string dimensions and

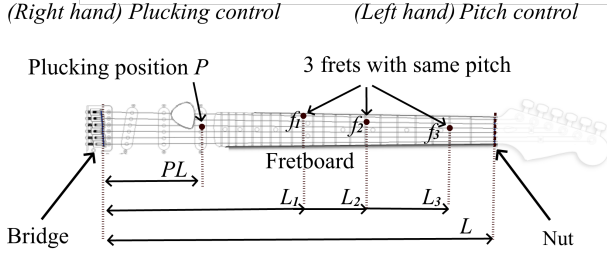


Figure 1: The plucking position is controlled by the right hand while the left hand controls pitch using the fretboard (adapted from [30]).

material properties, which are often available on string packets and in data sheets. The objective of the proposed method is to extract the location of interactions of both hands of the guitar player (as these can be arbitrarily located along a string) on a segment-by-segment basis from a single pitch guitar signal.

Figure 1 gives an overview of the guitar, where the right hand controls plucking position and the left hand controls pitch by using the fretboard. One pitch is produced in various positions and each string and fret combination is defined as one class, such that we have a set of  $K$  mutually exclusive classes. The feature set which we model in detail in the proposed simulation, consists of fundamental frequency  $\omega_0$  and inharmonicity  $B$  and it is extracted from the signal with the parametric pitch estimation method of [27]. This is contrary to other literature, where  $\omega_0$  and  $B$  are used as “black box” parameters, for accurate fundamental frequency estimation [28, 29] and feature extraction [5, 23, 26, 27].

## 2. SIGNAL MODEL

The string has an initial deflection  $\delta$  at the plucking position  $P$ , which is at the  $P$ th fraction of its length ( $0 < PL < L$ ), where  $L$  is the length of the vibrating string (see Fig. 1). Assuming that the string is inharmonic and pinned at locations  $l = 0$  and  $l = L$ , its general solution at time  $t \geq 0$  is a sum of  $M$  normal modes [17]

$$y(l, t) = \sum_{m=1}^M C_m \cos(\psi_m(\omega_0, B)t + \phi_m) \sin\left(\frac{m\pi l}{L}\right), \quad (1)$$

with fundamental frequency  $\omega_0$ , inharmonicity coefficient  $B$  and amplitude  $C_m$  and phase  $\phi_m$  of the  $m$ th mode. The instantaneous frequency  $\psi_m(\omega_0, B)$  follows the model  $\psi_m(\omega_0, B) = m\omega_0\sqrt{1 + Bm^2}$  [25]. Assuming the initial displacement to be triangular with no initial velocity  $\frac{\partial y}{\partial t} = 0 \forall l$ , the  $m$ th amplitude  $C_m$  is given by the Fourier sine series as [18, 19]

$$C_m = \frac{2\delta}{m^2\pi^2 P(1-P)} \sin(m\pi P). \quad (2)$$

From (2), it can be observed that the spectral envelope is sinusoidal along the partials, and dependent only on the plucking position  $P$  scaled by  $m^{-2}$  for the  $m$ th mode.

On the guitar, we assume that a pickup is measuring the displacement  $y(l, t)$  in close proximity to the string at location  $l = \lambda$ . At the discrete time instance  $n$  the observed signal  $x(n)$  is uniformly sampled such that is proportional to  $y$ , i.e.,  $x(n)|_{l=\lambda} \propto y(\lambda, t)$ . We parametrize  $x(n)$  as proposed in [27], where  $x(n)$  is modeled as an inharmonic sinusoidal part and a noise part  $v(n)$ ,

i.e.,

$$x(n) = \sum_{m=1}^M \alpha_m \exp(j\psi_m(\omega_0, B)n) + v(n), \quad (3)$$

where  $\alpha_m$  is the complex amplitude of the  $m$ th partial and  $M$  is the number of partials. At time instance  $n$ , the observed signal vector  $\mathbf{x} \in \mathbb{C}^N$  is represented as  $\mathbf{x} = [x(0) x(1) \cdots x(N-1)]^T$ , with  $T$  denoting the transpose. To simplify the notation,  $\psi_m(\omega_0, B)$  is denoted  $\psi_m$  in what follows. In matrix notation the observed signal is  $\mathbf{x} = \mathbf{Z}\boldsymbol{\alpha} + \mathbf{v}$ , where  $\boldsymbol{\alpha} = [\alpha_1 \cdots \alpha_M]^T$  is a vector containing complex amplitudes,  $\mathbf{v} = [v(0) v(1) \cdots v(N-1)]^T$  contains all noise terms and the Vandermonde matrix  $\mathbf{Z} \in \mathbb{C}^{N \times M}$  is defined as  $\mathbf{Z} = [\mathbf{z}(\psi_1) \mathbf{z}(\psi_2) \cdots \mathbf{z}(\psi_M)]$ , where each column is  $\mathbf{z}(\psi_m) = [1 e^{j\psi_m} e^{j\psi_m 2} \cdots e^{j\psi_m (N-1)}]^T$ . We denote the unknown and deterministic parameters that comprise the feature set  $\boldsymbol{\theta} = \{\omega_0, B, \boldsymbol{\alpha}\}$ . The amplitudes  $\boldsymbol{\alpha}$  are linear, while fundamental frequency  $\omega_0$  and inharmonicity  $B$  both are non-linear parameters.

## 3. PROPOSED METHOD

To obtain a class dependent model and to overcome the problem of the user having to provide audio recordings, as done in [12, 23, 26, 27], we propose a physical model of the feature space, only requiring knowledge of string properties. The proposed method implicitly increases computational efficiency on two levels; the user is not required to record audio for training, and the model is obtained from one simulation instead of features estimated from several audio recordings. Specifically, the features that we simulate are  $\omega_0$  and  $B$ , which have been proven useful for string and fret classification [23, 26, 27]. The string properties are:

- String length  $L$ , string tension  $T_0$ , plucking deflection  $\delta$ , core diameter  $d_c$ , core density  $\rho_c$ , wrapping diameter  $d_w$ , and wrapping density  $\rho_w$ .

We assume that the user knows  $L$  from the guitar-type, and the properties  $T_0$ ,  $d_c$ ,  $d_w$ ,  $\rho_c$  and  $\rho_w$  are available on string packaging or datasheets. The first feature  $\omega_0$  is well defined as

$$\omega_0 = \sqrt{\frac{T_0}{v}} \frac{1}{2L} \quad [\text{s}^{-1}], \quad (4)$$

where the mass per unit length  $v$  is found from string properties [31], i.e.,  $v = \frac{\pi}{4} \left( \rho_c d_c^2 + \rho_w [(2d_w + d_c)^2 - d_c^2] \right)$ . In the following, we extend on the state of the art [22] by providing a more accurate description of the inharmonicity coefficient. We hypothesize that this will provide more accurate string and fret classification.

### 3.1. Simulation of the inharmonicity coefficient

The authors of [21] state that string inharmonicity arises from two sources: the intrinsic inharmonicity due to the *stiffness* of the string and inharmonicity due to *stretching* caused by deflection. Assuming the definitions quantifying these effects as proposed in [22] can simply be added together, we propose the inharmonicity coefficient in the model of [25] to be:

$$B = \frac{\pi^3 E d_c^2}{16 T_0 L^2} \left( \frac{d_c^2}{4} + \frac{3\delta^2}{8} \right) = \frac{\pi^3 E d_c^2 (2d_c^2 + 3\delta^2)}{128 T_0 L^2} \quad [\cdot]. \quad (5)$$

where Young's Modulus  $E$  of the string material is defined in (8). It is expected that  $d_c \ll \delta$ , i.e., the inharmonicity caused by the plucking deflection dominates the intrinsic inharmonicity. However, this

is only true immediately after the plucking event, as the effect of  $\delta$  dies down rapidly after the string is excited [22]. In our application, we solely use the first 40 ms of the plucked string sound, as we hypothesize that it is of importance to include the effect that the initial displacement has on the total inharmonicity. In [22], for the use of  $\delta$  it is assumed that  $P = 0.5$ , i.e. the string is plucked at the center. Realistically, this is rarely the case, which is why we propose to calculate  $\delta$  in (5) from the transverse displacement at any  $P$ th fraction of the string-length. We define the change in string length due to plucking  $\Delta L_P$  as

$$\Delta L_P = \sqrt{(LP)^2 + \delta_P^2} + \sqrt{(L(1-P))^2 + \delta_P^2} - L \quad [\text{m}]. \quad (6)$$

This will also retain the tension present in the string as  $T_0$  and  $\Delta L$  are proportional according to (8). By setting  $P = 0.5$ , we simplify (6) such that  $\Delta L_{0.5} = 2\sqrt{(0.5L)^2 + \delta^2} - L$ , after which  $\Delta L_{0.5}$  is replaced with the length extension  $\Delta L_P$  in (6) and solving for  $\delta$  yields

$$\delta = \frac{\sqrt{\Delta L_P^2 + 2L\Delta L_P}}{2} \quad [\text{m}], \quad (7)$$

which can be used to calculate the inharmonicity due to plucking for any  $P$ . When the string is brought to pitch (i.e., tuned) it stretches according to its Young's modulus (or elastic modulus)  $E$  which is defined as

$$E = \frac{\text{Tensile stress}}{\text{Strain}} = \frac{T_0/A_c}{\Delta L/(L - \Delta L)} \quad [\text{N} \cdot \text{m}^{-2}], \quad (8)$$

where  $A_c = \pi d_c^2/4$  is the cross-sectional area of the string-core and  $\Delta L$  is the amount by which the string is stretched. The variables  $E$ ,  $A_c$  and  $L$  are often considered constant and we know that  $T_0$  and  $\Delta L$  are proportional. It is important to note that we use  $L - \Delta L$  to describe the original string length before stretched to  $L$ . The Young's modulus is a property defined for solid materials. Even though the string core and wrapping could be the same material, the wrapping is not solid. This means that when calculating the total inharmonicity using (5) for a wrapped string, the Young's modulus – or more specifically,  $T_0$  – will be incorrect. To overcome this problem we express the total tension as  $T_0 = T_w + T_c$ , where  $T_w$  is the tension contributed by the wrapping and  $T_c$  the tension contributed by the core. We can assume the wrapping to be a spring as in [32] having the spring constant  $k = \frac{Gd_w^4}{8ND^3}$  [33]. Here,  $G$  is the shear modulus of the wrapping material,  $N = \frac{L - \Delta L}{d_w}$  is the number of turns in the coil (assuming that for the untensioned string, every coil touches the previous and the next coil) and  $D = d_w + d_c$  is the mean diameter of the spring [32]. The length extension  $\Delta L$  for a spring is

$$\Delta L = \frac{T_w}{k} = \frac{8T_wLD^3}{Gd_w^5 + 8T_wD^3} \quad [\text{m}]. \quad (9)$$

For the core of the string we can rewrite (8), i.e.,

$$\Delta L = \frac{LT_c}{A_cE + T_c}, \quad [\text{m}], \quad (10)$$

and as the core and wrapping are part of the same string, they undergo the same length-extension  $\Delta L$ , meaning that we can set (9) and (10) equal to each other and (recalling  $T_0 = T_w + T_c$ ) we can solve for  $T_w$  and  $T_c$

$$T_w = \frac{GT_0d_w^5}{Gd_w^5 + 8A_cD^3E}, \quad T_c = \frac{8A_cD^3ET_0}{Gd_w^5 + 8A_cD^3E} \quad [\text{N}]. \quad (11)$$

Finally, inserting  $T_w$  into (9) or  $T_c$  into (10), gives the length extension based on the string-material properties of both core and wrapping, which can then be used in (8) to calculate the correct Young's modulus  $E$ .

This concludes the derivations used for the proposed method of simulating the feature set  $\{\omega_0, B\}$ . In the following we explain how to estimate these from the observed signal  $\mathbf{x}$  and how to classify  $\mathbf{x}$  as being produced from a specific string and fret combination.

### 3.2. Estimation of String, Fret and Plucking Position

The feature set  $\theta$  is estimated using the principle of maximum likelihood. In doing so, the observed signal distribution is modeled as circular complex white Gaussian noise, and as derived in [27, 34], a computationally efficient approach can be shown to be

$$\{\hat{\omega}_0, \hat{B}\} = \underset{\omega_0, B}{\operatorname{argmax}} \left\| \mathbf{Z}^H \mathbf{x} \right\|_2^2, \quad (12)$$

which can be implemented using just one FFT per segment, i.e.,  $\|\mathbf{Z}^H \mathbf{x}\|_2^2 = \sum_{m=1}^M |X(\psi_m(\omega_0, B))|^2$  with  $X(\cdot)$  being the FFT of  $\mathbf{x}$ . To increase computational efficiency, an initial fundamental frequency estimate is obtained with  $B = 0$ , and then a two dimensional search grid is defined for estimating  $\{\hat{\omega}_0, \hat{B}\}$ . Finally, the amplitudes of the inharmonic partials are found using least squares, i.e.,  $\hat{\alpha} = (\mathbf{Z}^H \mathbf{Z})^{-1} \mathbf{Z}^H \mathbf{x}$ . Having found the fundamental frequency and the inharmonicity parameters  $\phi = [\hat{\omega}_0, \hat{B}]^T$  from the observed signal snapshot  $\mathbf{x}$ , the next step is to classify the observed signal as being produced from a string and fret combination.

We have a set of  $K$  mutually exclusive classes  $\Gamma = \{\gamma_1, \dots, \gamma_K\}$  representing all possible string and fret combinations. The MAP classifier with decision function  $\hat{\gamma}(\cdot) : \mathbb{R}^I \rightarrow \Gamma$  is  $\hat{\gamma}(\phi) = \underset{\gamma \in \Gamma}{\operatorname{argmax}} p(\phi|\gamma)P(\gamma)$ . We model  $\phi$  as coming from a normal object with class  $\gamma_k$ , then the  $k$ th conditional probability density is  $p(\phi|\gamma_k) = \mathcal{N}(\mu_k, \Lambda_k)$  where the expectation vector  $\mu_k$  and covariance matrix  $\Lambda_k$  are computed from a Monte Carlo simulation using (4) and (5), where we assume all physical properties of the strings are normally distributed random variables. The covariance matrices are here modeled as being class dependent, unlike the state of the art [27]. Neglecting terms that do not depend on the class index  $k$  yields the following, classification scheme  $\hat{\gamma}(\phi) = \gamma_i$ , with

$$i = \underset{k=1, \dots, K}{\operatorname{argmax}} \left\{ -\ln|\Lambda_k| + 2\ln P(\gamma_k) - (\phi - \mu_k)^T \Lambda_k^{-1} (\phi - \mu_k) \right\}. \quad (13)$$

The prior  $P(\gamma_k)$  can be specified from the number of training samples from class  $\gamma_k$  or simply be assumed uniform.

As demonstrated in [27], once the partial amplitudes have been estimated using  $\hat{\alpha} = (\mathbf{Z}^H \mathbf{Z})^{-1} \mathbf{Z}^H \mathbf{x}$ , the plucking position  $\hat{P}$  can be found. As proposed in [27] we minimize the log spectral (LS) distance between the estimated amplitudes  $\hat{\alpha}$  and the model  $\mathbf{C}$  to find  $\hat{P}$ , i.e.,

$$\hat{P} = \underset{P}{\operatorname{argmin}} (d_{LS}(\hat{\alpha}, \mathbf{C})), \quad (14)$$

where  $\mathbf{C}$  is obtained from the model in (2). To emphasize the dependency on plucking position  $\mathbf{C}$  can be written as a function of  $P$ , i.e.,  $\mathbf{C}(P) = [C_1(P), C_2(P), \dots, C_M(P)]^T$  and the log spectral distortion is thus defined as  $d_{LS}(\hat{\alpha}, \mathbf{C}(P)) = \sqrt{\frac{1}{M} \sum_m 10 \log_{10} \frac{|\hat{\alpha}_m|^2}{|C_m(P)|^2}}$ .

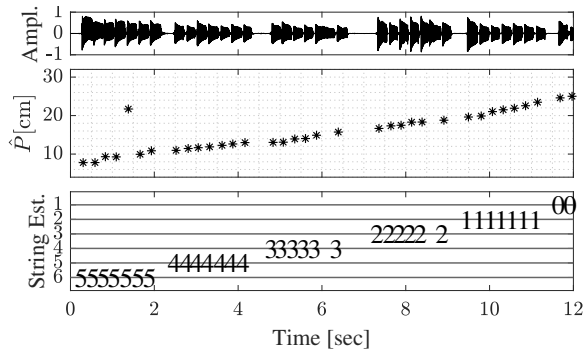


Figure 2: String, fret and plucking position estimates with moving plucking position and moving string and fret for electric guitar.

#### 4. EVALUATION AND DISCUSSION

This section describes the experiments conducted to evaluate the proposed string and fret classifier along with the plucking position estimator. These experiments aim to demonstrate that the physical model helps to obtain a class dependent model for the classifier. They also aim to show that it is possible to classify string and fret from the simulated feature space, and to determine to what extent the classifier is robust towards noise. Lastly, we demonstrate an application of detailed transcription of string, fret and plucking position. The classifier is compared to [27] and tested on 1560 recordings of two guitars from [27], namely an electric Les Paul Firebrand with Elixir Nanoweb (.010-.054) strings and an acoustic Martin DR with SP (.012-.052) strings. The data and MATLAB code are available online<sup>1</sup> and we refer to the available code for implementation details. The estimation and classification is done from one 40 ms segment for each recording, extracted at the onset event using [35].

Figure 2 demonstrates an example application of detailed transcription, where we can observe that the proposed method is capable of estimating string, fret and plucking position, when the electric guitar is played fast with changing string and fret combinations and a moving plucking position, from the bridge towards the nut. Figure 3 shows the sensitivity of the classifier to noise. Here, all observed signals  $\mathbf{x}$  have been induced with additive white Gaussian noise (AWGN) at various levels of average signal to noise ratio (SNR). A plausibility filter as the one proposed in [12] can be applied to the classifier such that only a few candidates out of all  $K$  classes are probable, hence speeding up the algorithm. The state-of-the-art model [27] is trained on 60 recordings obtained from the 12th fret on all six strings. It can be observed that the proposed method performs slightly better than the trained model of [27]. Interestingly, we can see from this graph that the method has a higher error-rate at a higher SNR for the electric guitar than for the acoustic, which implies that the higher harmonics of the latter have a relatively higher amount of energy than in the electric signal, which can be related to the electronics of the transducers on these guitars.

Finally, we can evaluate the detailed and class dependent simulation of the feature space and compare it to estimates obtained with [27]. In this simulation, all string properties are normally distributed with a standard deviation specified as 0.5 % of their respective mean value (see the available evaluation code). The resulting distributions are shown in Figure 4, where the consistency between simulation and measurement can clearly be observed. Note here,

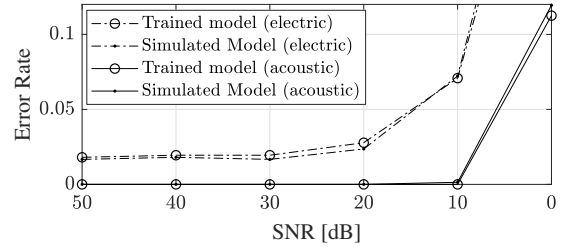


Figure 3: String, fret classification error rate as a function of SNR. Each marker represents 720 classifications. The trained model represents [27] and the simulated model is the proposed method.

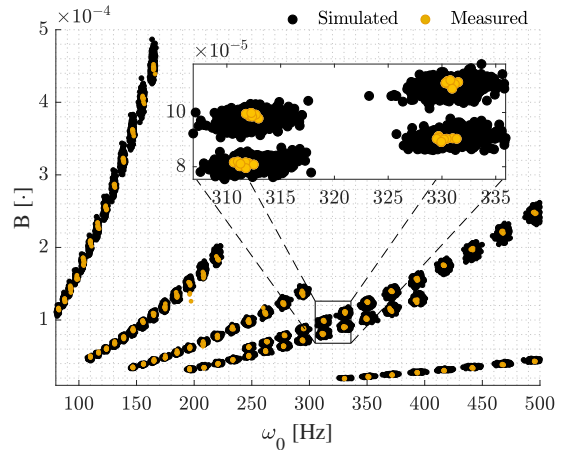


Figure 4: Comparison of simulated random variables to the measured estimates of the features for the acoustic guitar. For each class there are 500 simulated values and 10 estimated measurements.

that the simulation obtains a detailed and class dependent distribution of  $\omega_0$  and  $B$ , which is not possible to obtain with state-of-the-art method [27]. In the proposed simulation framework, we have not observed that a plucking deflection influences the inharmonicity as denoted by (5). As stated in [22], the effect of  $\delta$  dies down rapidly after the string excitation, and we consider this to happen in a time-period an order smaller than 40 ms. We also emphasize that the computation time for simulating a model with 500 realizations of 78 classes takes 45 ms on an i7 processor using MATLAB.

#### 5. CONCLUSION

In this paper, a fast and effective method for estimation of guitar string, fret and plucking position based on a physical model of string excitation and vibration was proposed. The proposed classifier uses physical properties of the guitar string to simulate a class dependent pitch and inharmonicity distribution. The classifier was proven to perform as good as [27] under noisy conditions, and since no audio needs to be recorded to build a model, it is faster than state-of-the-art methods by nature. The plucking position estimator is the minimizer of the log spectral distance between the amplitudes of the observed signal and the plucking model and it is evaluated in a proof-of-concept experiment with sudden changes of string, fret and plucking positions, with accurate results. The estimator operates on a 40 ms segment-by-segment basis and is suitable for high tempo and real-time applications.

<sup>1</sup><https://tinyurl.com/waspaa19>

## 6. REFERENCES

- [1] C. R. Sullivan, "Extending the Karplus-Strong algorithm to synthesize electric guitar timbres with distortion and feedback," *Computer Music Journal*, vol. 14, no. 3, pp. 26–37, 1990.
- [2] M. Karjalainen, V. Välimäki, and Z. Jánosy, "Towards high-quality sound synthesis of the guitar and string instruments," *Proc. Int. Computer Music Conf.*, pp. 56–63, 1993.
- [3] M. Laurson, C. Erkut, V. Välimäki, and M. Kuuskankare, "Methods for modeling realistic playing in acoustic guitar synthesis," *Computer Music Journal*, vol. 25, no. 3, pp. 38–49, 2001.
- [4] C. Kehling, J. Abeßer, C. Dittmar, and G. Schuller, "Automatic tablature transcription of electric guitar recordings by estimation of score- and instrument-related parameters," *Proc. Int. Conf. Digital Audio Effects*, 2014.
- [5] J. Abeßer and H. Lukashovich, "Feature-based extraction of plucking and expression styles of the electric bass," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2012.
- [6] A. Paté, J.-L. Le Carrou, and B. Fabre, "Predicting the decay time of solid body electric guitar tones," *J. Acoust. Soc. Am.*, vol. 135, no. 5, pp. 3045–3055, 2014.
- [7] Z. Mohamad, S. Dixon, and C. Harte, "Pickup position and plucking point estimation on an electric guitar," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2017.
- [8] C. Traube and J. O. Smith, "Estimating the plucking point on a guitar string," *Proc. Int. Conf. Digital Audio Effects*, 2000.
- [9] H. Penttinen and V. Välimäki, "A time-domain approach to estimating the plucking point of guitar tones obtained with an under-saddle pickup," *Applied Acoustics*, vol. 65, no. 12, pp. 1207–1220, 2004.
- [10] H. Penttinen, J. Siiskonen, and V. Välimäki, "Acoustic guitar plucking point estimation in real time," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 3, pp. 209–212, 2005.
- [11] C. Erkut, V. Välimäki, M. Karjalainen, and M. Laurson, "Extraction of physical and expressive parameters for model-based sound synthesis of the classical guitar," in *Audio Eng. Soc.*, 2000.
- [12] J. Abeßer, "Automatic string detection for bass guitar and electric guitar," *From Sounds to Music and Emotions*, pp. 333–352, 2013.
- [13] F. Germain and G. Evangelista, "Synthesis of guitar by digital waveguides: Modeling the plectrum in the physical interaction of the player with the instrument," *Proc. IEEE Workshop on Appl. of Signal Process. to Aud. and Acoust.*, pp. 25–28, 2009.
- [14] G. Evangelista and F. Eckerholm, "Player-instrument interaction models for digital waveguide synthesis of guitar: Touch and collisions," *IEEE Trans. Audio, Speech, Language Process.*, vol. 18, no. 4, pp. 822–832, 2010.
- [15] S. Poirot, S. Bilbao, M. Aramaki, and R. Kronland Martinet, "Sound morphologies due to non-linear interactions: Towards a perceptive control of environmental sound-synthesis processes," in *Proc. Int. Conf. Digital Audio Effects*, 2018.
- [16] S. Bilbao and A. Torin, "Numerical modeling and sound synthesis for articulated string/fretboard interactions," *J. Audio Eng. Soc.*, vol. 63, no. 5, pp. 336–347, 2015.
- [17] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, 2nd ed. Springer, 1998.
- [18] —, *Principles of vibration and sound*, 2nd ed. Springer, 2004.
- [19] W. F. Donkin, *Acoustics*, 2nd ed. Oxford Clarendon Press, 1884.
- [20] L. Rayleigh, *The Theory Of Sound*, 1st ed. MacMillan and Co. Ltd. London, 1894.
- [21] R. S. Shankland and J. W. Coltman, "The departure of the overtones of a vibrating wire from a true harmonic series," *J. Acoust. Soc. Am.*, vol. 10, no. 3, pp. 161–166, 1939.
- [22] T. D. Rossing, *The science of string instruments*, 1st ed. Springer, 2010.
- [23] I. Barbancho, L. Tardon, S. Sammartino, and A. Barbancho, "Inharmonicity-based method for the automatic generation of guitar tablature," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 6, pp. 1857–1868, 2012.
- [24] C. Dittmar, A. Männchen, and J. Abeßer, "Real-time guitar string detection for music education software," *Proc. Int. Workshop on Image Analysis for Multimedia Interactive Services*, pp. 1–4, 2013.
- [25] H. Fletcher, "Normal vibration frequencies of a stiff piano string," *J. Acoust. Soc. Am.*, vol. 36, pp. 203–209, 1964.
- [26] J. J. Michelson, R. M. Stern, and T. M. Sullivan, "Automatic guitar tablature transcription from audio using inharmonicity regression and Bayesian classification," in *Audio Eng. Soc.*, 2018.
- [27] J. Hjerrild and M. Christensen, "Estimation of guitar string, fret and plucking position using parametric pitch estimation," *Accepted for Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2019. [Online]. Available: [http://vbn.aau.dk/files/296639129/icassp-plucking\\_position\\_estimation.pdf](http://vbn.aau.dk/files/296639129/icassp-plucking_position_estimation.pdf)
- [28] A. Galembo and A. Askenfelt, "Measuring inharmonicity through pitch extraction," *Journal STL-QPSR*, vol. 35, no. 1, pp. 135–144, 1994.
- [29] T. Nilsson, S. Adalbjörnsson, N. Butt, and A. Jakobsen, "Multi-pitch estimation of inharmonic signals," *Proc. European Signal Processing Conf.*, pp. 1–5, 2013.
- [30] J. M. Phillips, "Jeffrey phillips design," 2018. [Online]. Available: <http://www.imjeffp.com/>
- [31] I. Firth, R. Bain, and A. Gallaher, "String design equations," *Journal of the Catgut Acoustical Society*, vol. 46, pp. 3–6, 1984.
- [32] J. A. Kemp, "The physics of unwound and wound strings on the electric guitar applied to the pitch intervals produced by tremolo/vibrato arm systems," *PLOS One*, 2017.
- [33] P. R. N. Childs, *Mechanical Design Engineering Handbook*. Butterworth-Heinemann, 2014.
- [34] M. Christensen and A. Jakobsen, *Multi-Pitch Estimation*, 1st ed. Morgan and Claypool, 2009.
- [35] O. Lartillot and P. Toiviainen, "A Matlab toolbox for musical feature extraction," *Proc. Int. Conf. Digital Audio Effects*, 2000.